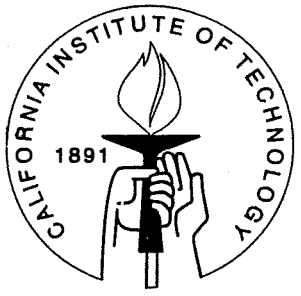
DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

In or Out?: Centralization by Majority Vote

Jacques Crémer
CNRS, GREMAQ and IDEI, Université de Toulouse

Thomas R. Palfrey
California Insitute of Technology



SOCIAL SCIENCE WORKING PAPER 879

April 1994

In or Out?: Centralization by Majority Vote

Jacques Crémer Thomas R. Palfrey

Abstract

We present a positive theory of centralization of political decisions. Voters choose centralization or decentralization depending on their forecast of the political organization that will favor the policies they prefer. We study the induced preferences for centralization as well results of different forms of referenda.

Résumé

Nous présentons une théorie positive de la centralisation des décisions politiques. Les électeurs choisissent la forme d'organisation des services publics qui favorisent les politiques qu'ils préfèrent. Nous étudions les préférences induites pour la centralisation et comparons les résultats de différentes formes de consultations populaires.

Keywords: voting, unification, decentralization

JEL Classification number: 025, 320, 940

In or Out?: Centralization by Majority Vote

Jacques Crémer Thomas R. Palfrey

1 Introduction

Political bodies often must choose whether or not to adopt a single centralized policy on an issue. This paper investigates some of the factors that determine how this *centralization question* is decided. We focus on a single but important dimension of this question, the tradeoffs faced by a collection of small component political entities (which we will call *districts*) between maintaining a sovereign right to shape policies such as education, the environment, or social conduct, specifically tailored to local needs and preferences, versus the alternative of agreeing to follow uniform (across districts) policies along each of these different dimensions. We call such a committment to policy uniformity across a subset of districts *centralization.*[1]

Contrary to most of the economic iterature on this topic our approach is positive. We do not discuss the efficiency of centralization, but focus exclusively on the "preference aggregation" problem inherent in centralization. Therefore in our model, all outcomes are Pareto optimal, but the policy will depend on the level of centralization. In the centralized case all districts will have the same policy. In the totally decentralized case each different district will adopt a different policy, and the distribution of these policies, adopted under majority rule can vary considerably. Voters will vote for the allocation of responsibilities to the level of government which is the most likely to choose the policy they prefer. Under these conditions, our aim is to predict the outcome of a referendum on centralization. In order to do so we will have to study the induced preferences for or against centralization, which depend on the distribution of policy preferences across all districts.

The model excludes from consideration efficiency issues relating to centralization, such as economies (or diseconomies) of scale in the production of public goods and services, networking externalities, benefits from coordination and standardization.[2] These are

[1] This is closely related theoretically to the problem of confederation of several states under a unified constitution. The results carry over to that setting.

[2] For recent discussions of some of these normative issues see Oates [1972], Sinn [1993], Cremer, Estache and Seabright [1993] and the references they cite.

important factors which often seem to favor centralization.[3] We exclude them from our model because they would confound the pure "preference aggregation" effect on induced preferences for centralization.

In our analysis, districts and unions are composed of individuals; these individuals have diverse preferences over the various policy issues that are decided by governments; and there is some correlation between the tastes of individuals that belong to the same district. Consequently, the outcomes and policies decided at the decentralized level vary widely across the set of feasible alternative policies. These diverse preferences and wide range of decentralized outcomes imply that some individuals will be better off committing in advance to adhere to uniform policies, while others would prefer not to. The impetus for centralization will be greater, and the pressures for fragmentation will be less, the larger the proportion of individuals that prefer committment to uniform policies.

In our model, the main benefit of centralization is *policy moderation*: the policies of the union are close to the average of the policies of the districts. In addition, moderation will manifest itself in the fact that it is easier to predict a centralized policy than the policy of a district. At one extreme, the centralized policy is virtually certain to be at the ex ante median of the whole population of voters, while there is some variation across district medians. More generally, there will be less uncertainty about the union median than in the district medians. This policy moderation effect plays a key role in our analysis.

In this way we obtain derived preferences for centralization as a function of a voter's type, which is determined by his ideal policy and the district to which he belongs. These derived preferences have a very simple structure. Voters with extreme ideal policies prefer decentralization; voters with moderate ideal policies prefer centralization. Voters are also more likely to support centralization the larger the district to which they belong, and the closer to uniform the distribution of sizes of the other districts.

Having characterized derived preferences, we will be able to study the relationship between the distribution of voter preferences and the likelihood of centralization. The answer depends on the level at which votes are aggregated. If votes are aggregated at the union level, centralization vote will rarely succeed. The expected percentage of "YES" votes for centralization is bounded above at only slightly more than fifty percent. Centralization is more likely when votes are aggregated at the district level, which is the institution more favored in the real world.[4] The expected percentage of districts with majorities favoring centralization is greater than the expected percentage of the overall population favoring centralization, unless there is so much heterogeneity across districts that centralization has very little (expected) support. We call this the *principle of aggregation.*

[3]This is not exclusively true–for example there may be inefficiencies associated with establishing a policy "monopoly" and there do exist problems of control of politicians that might be best resolved by decentralization.

[4]For example, in the approval and amendment of national constitutions, and major international agreements such as the European agreement at Maastricht.

In the next section, we describe the model and characterize induced preferences for centralization. We also conduct some comparative statics on the number of districts and their size distribution. Section 3 compares voting procedures. We study the consequences of the method used to aggregate votes, and explains how the principle of aggregation applies to our model. Then, we explore the effect of an "opt out" provision that allows districts that vote against centralization to adopt their own separate policies. Section 4 presents some extensions, and section 5 some concluding remarks.

2 Basic results

2.1 The model

A *federation* is composed of a finite set $\mathcal{D}$ of *districts*. In each district, there is a continuum of individuals. For most of this paper, there is only one issue on which a policy is to be chosen, and the issue space is represented by the real line.

Each individual has an *ideal policy* or *point* t, which we will also call his *type*. If policy x is adopted the utility of an individual of type t is

$$U(x;t) = -(x-t)^2.$$

When voting over constitutions, the agent will be uncertain about the outcome. In this case, between two distributions over policies he will choose that distribution that maximizes his expected utility.

An individual is represented by his type t and the district d to which he belongs. Within district d, the types of the individuals are normally distributed with mean m_d and variance 1 (this normalization is harmless). The district means are normal independent random variables, normally distributed with mean 0 and variance σ_d^2. Each individual only knows his type, the district to which he belongs, and the statistical distribution of preferences.

The median type of the whole federation is[5]

$$M = \sum_{d\in\mathcal{D}} \alpha_d m_d,$$

where $\alpha_d \in (0,1)$ is the relative size of district d:

$$\sum_{d\in\mathcal{D}} \alpha_d = 1.$$

[5]Since each district has a continuum of voters, in this model the median is equal to the mean.

We assume that the political process within any district, in the case of decentralization, or in the federation taken as a whole, in the case of centralization, yields a policy that is the ideal point of the median voter. Before a vote is taken on the policy, a vote on the centralization question is first taken. If centralization wins, then a single policy is taken for the whole federation. If centralization fails, each district adopts their median voter's ideal point.

Individual (t, d) votes for centralization if and only if his disutility under decentralization is less than his disutility under centralization, that is if

$$E[(t - m_d)^2 | t, d] \leq E[(t - M)^2 | t, d]. \tag{1}$$

Note that in general an individual's estimate of the mean of all types in the population will depend not only on his own type, but also of the district to which he belongs (more precisely on its size).

2.2 Who votes for decentralization?

Let us first compute the left hand side of equation (1). Using standard results from De Groot (1970), chapter 5, one can easily prove that the distribution of m_d conditional on t is normal with mean

$$\frac{\sigma_d^2}{\sigma_d^2 + 1} t$$

and variance

$$\frac{\sigma_d^2}{\sigma_d^2 + 1}.$$

This implies

$$\begin{aligned}
E[(m_d - t)^2 | t, d] &= E\left[((m_d - E[m_d|t,d]) + (E[m_d|t,d] - t))^2 \, | t, d\right] \\
&= E\left[\left((m_d - E[m_d|t,d]) + (\sigma_d^2/(1+\sigma_d^2) - 1)t\right)^2 | t, d\right] \\
&= \text{var}\,[m_d|t,d] + \left(\frac{1}{\sigma_d^2 + 1}\right)^2 t^2 \\
&= \frac{\sigma_d^2}{\sigma_d^2 + 1} + \left(\frac{1}{\sigma_d^2 + 1}\right)^2 t^2.
\end{aligned}$$

To compute the right hand side of equation (1) note first that the distribution of M conditional on t and d is Normal with mean

$$\mathrm{E}[M|t,d] = (1-\alpha_d)\cdot 0 + \alpha_d \frac{\sigma_d^2}{\sigma_d^2+1} t = \frac{\alpha_d \sigma_d^2}{1+\sigma_d^2} t,$$

and variance:

$$\begin{aligned} \mathrm{var}[M|t,d] &= \sum_{i\in\mathcal{D}} \alpha_i^2 \sigma_d^2 - \alpha_d^2 \frac{\sigma_d^4}{\sigma_d^2+1} \\ &= \frac{\sigma_d^4(\mu-\alpha_d^2)+\sigma_d^2\mu}{1+\sigma_d^2}, \end{aligned}$$

where

$$\mu = \sum_{d\in\mathcal{D}} \alpha_d^2.$$

Therefore, the right hand side of (1) is equal to

$$\begin{aligned} \mathrm{E}\left[(M-t)^2|t,d\right] &= \mathrm{E}\left[(M-\mathrm{E}[M|t,d])^2 + (\mathrm{E}[M|t,d]-t)^2|t,d\right] \\ &= \mathrm{var}[M|t,d] + (\mathrm{E}[M|t,d]-t)^2 \\ &= \frac{\sigma_d^4(\mu-\alpha_d^2)+\sigma_d^2\mu}{1+\sigma_d^2} + \frac{[(1-\alpha_d)\sigma_d^2+1]^2}{(1+\sigma_d^2)^2}. \end{aligned}$$

We can therefore rewrite equation (1) under the form

$$\frac{\sigma_d^2}{\sigma_d^2+1} + \left(\frac{1}{\sigma_d^2+1}\right)^2 t^2 \leq \frac{\sigma_d^4(\mu-\alpha_d^2)+\sigma_d^2\mu}{1+\sigma_d^2} + \frac{[(1-\alpha_d)\sigma_d^2+1]^2}{(1+\sigma_d^2)^2},$$

which can be rewritten as:

$$\frac{t^2}{\sigma_d^2+1} \leq \frac{1-\mu-\sigma_d^2(\mu-\alpha_d^2)}{(1-\alpha_d)((1-\alpha_d)\sigma_d^2+2)}. \tag{2}$$

The denominator of the left hand side of (2) is the unconditional variance of the distribution of the type of an agent. Let us denote by Φ the cumulative distribution function of the unit Normal. Equation (2) implies that the expected proportion of agents who favor decentralization in district d is

$$2\left[1-\Phi\left(\sqrt{\frac{1-\mu-\sigma_d^2(\mu-\alpha_d^2)}{(1-\alpha_d)((1-\alpha_d)\sigma_d^2+2)}}\right)\right],$$

when the right term is well defined. When it is not well defined, i.e., when the right hand side of equation (2) is negative, all agents in the district vote for centralization.

Voters trade off the expected closeness of the median of their own district against the predictable moderation of the federation median, which is unlikely to be very far from the most moderate policy 0. Therefore, extreme voters will tend to prefer decentralized outcomes (since moderation is bad from their point of view) whereas sufficiently moderate voters will value the predictable moderation of the federation median.

Voters in large districts have better information about the federation median than voters in small districts. Moreover, a voter in a large district expects the median of the federation to be closer to his ideal point than a voter with identical preferences in a smaller district, since the large district median has a greater weight in determining the federation median. Formally:

Proposition 1 *For all t and for all d, d' such that $\alpha_d > \alpha_d'$, if (t, d') weakly prefers centralization then (t, d) strictly prefers centralization.*

Proof: The denominator of equation (2) is strictly decreasing in α_d and its numerator is strictly increasing in α_d. The result follows. ∎

We also have the following proposition:

Proposition 2 *If*

$$\alpha_d^2 \leq \mu - \frac{1-\mu}{\sigma_d^2}$$

all agents in district d prefer decentralization.

Proof: The denominator of the right hand side of equation (2) is positive. Its numerator is less than or equal to 0 if and only if the inequality in this proposition is satisfied. Since all agents prefer decentralization if and only if the right hand side of (2) is less than or equal to 0, the result follows. ∎

2.3 The effect of the size distribution of districts

In order to generate more insights on the determinants of votes for centralization and decentralization, it is useful to examine a few special cases.

2.3.1 Equal sized districts

Assume that the set $\mathcal{D}$ is composed of N districts with the same population. Then $\alpha_d = 1/N$ for all d, and equation (2) reduces to

$$\begin{aligned} \frac{t^2}{\sigma_d^2+1} &\leq \frac{1-\sigma_d^2/N}{2+(N-1)\sigma_d^2/N} \\ &= 1 - \frac{1+\sigma_d^2}{2+\sigma_d^2-\sigma_d^2/N}, \end{aligned}$$

As N increases, for fixed σ_d^2, the right hand side of (2.3.1) increases. Hence, with equal-sized districts, an increase in the number of districts favors centralization. On the surface, this would seem to contradict Proposition 1, which states that it is the agents that belong to larger districts that are more likely to favor centralization. However, that proposition holds the distribution of α_d's constant, whereas in the present analysis this distribution changes as N changes.

As N converges to infinity the right hand side of (2.3.1) converges to

$$\frac{1}{2+\sigma_d^2}.$$

The comparative statics are reasonable. The proportion of individuals favoring decentralization increases with σ_d^2. If σ_d^2 is large enough, all agents prefer decentralization since the district variance swamps the idiosyncratic variance. As σ_d^2 converges to 0, the proportion of agents who favor centralization converges from below to $2\left(1-\Phi\left(\frac{1}{\sqrt{2}}\right)\right) \approx 0.52$.

2.3.2 A dominant district and a fringe of many small districts

With one very large district and many small ones of equal sizes, μ will be approximately equal to the square of the population share of the large district. The proportion of the population of the small districts that will vote for decentralization is

$$2\left[1-\Phi\left(\sqrt{\frac{1-\mu-(1+\sigma_d^2)}{2+\sigma_d^2}}\right)\right], \tag{3}$$

and the proportion of the population of the large district that will vote for decentralization is

$$2\left[1-\Phi\left(\sqrt{\frac{1}{2+(1-\mu)\sigma_d^2}}\right)\right] \tag{4}$$

From proposition 1, we know that the proportion of the population that favors decentralization will be larger in the small districts. From (3) and (4), as μ increases a larger proportion of the population of small districts and a smaller proportion of the population of the large district votes for decentralization. However, even if less than 30 percent of the population belongs to the large district, at least 50 percent of the population of the small districts will vote for centralization, whatever σ_d^2. Indeed, for any $\sigma_d^2 > 0$

$$2\left[1-\Phi\left(\sqrt{\frac{1-\mu-(1+\sigma_d^2)}{2+\sigma_d^2}}\right)\right] < 2\left[1-\Phi\left(\sqrt{\frac{1-\mu}{2}}\right)\right],$$

and the right hand side of this inequality is equal to .5 when μ is approximately equal to $(.30)^2$.

These comparative statics make sense: not only are large districts more favorable to centralization, but small districts are more willing to centralize policy decisions with them the less dominant they are.

3 Comparing voting procedures

3.1 Aggregation of votes at the district level

Up to this point, the discussion has focussed on the preferences of agents for centralization or decentralization. We have not described carefully the way in which votes were translated into a centralization decision. In this subsection we compare two types of procedures:

- Under *federated referendum procedures*, votes are aggregated at the district level, and centralization prevails if and only if it is favored by the majority of districts;
- in contrast in *general referendum procedures* whichever centralization decision favored by the greater number of voters in the aggregate is chosen.

Federated referendum procedures are common practice in Western democracies when amending the constitution, for example. Representative institutions, such as parliaments

and congress, also typically aggregate preferences at the district level by electing representatives. Then, a centralized policy is enacted by legislation if a majority of the district representatives approves.

Our aim is to identify which of these two types of referendum procedures is more favorable to centralization under the assumption that there is a very large number of equal size districts, as in 2.3.1 with N set equal to $+\infty$. In this case $\alpha_d = 0$ for all d and furthermore $\mu = 0$. From equation (2) individuals whose types satisfy

$$t^2 \leq \frac{1+\sigma_d^2}{2+\sigma_d^2}$$

will vote for centralization. Let

$$t^*(\sigma_d) \stackrel{\text{def}}{=} \sqrt{\frac{1+\sigma_d^2}{2+\sigma_d^2}}.$$

The proportion of individuals of district d whose types lie in the interval $[-t^*(\sigma_d),\ +t^*(\sigma_d)]$ is

$$\Phi(t^*(\sigma_d) - m_d) - \Phi(-t^*(\sigma_d) - m_d) = \Phi(t^*(\sigma_d) - m_d) + \Phi(t^*(\sigma_d) + m_d) - 1.$$

The district will approve centralization if this proportion is at least equal to $1/2$, that is if

$$\Phi(t^*(\sigma_d) - m_d) + \Phi(t^*(\sigma_d) + m_d) \geq 3/2. \tag{5}$$

This expression defines a limit value $\widetilde{m}(\sigma_d)$ such that a district will vote for centralization if and only if its median lies in the interval $[-\widetilde{m}(\sigma_d), +\widetilde{m}(\sigma_d)]$, and the proportion of districts that approve centralization is therefore

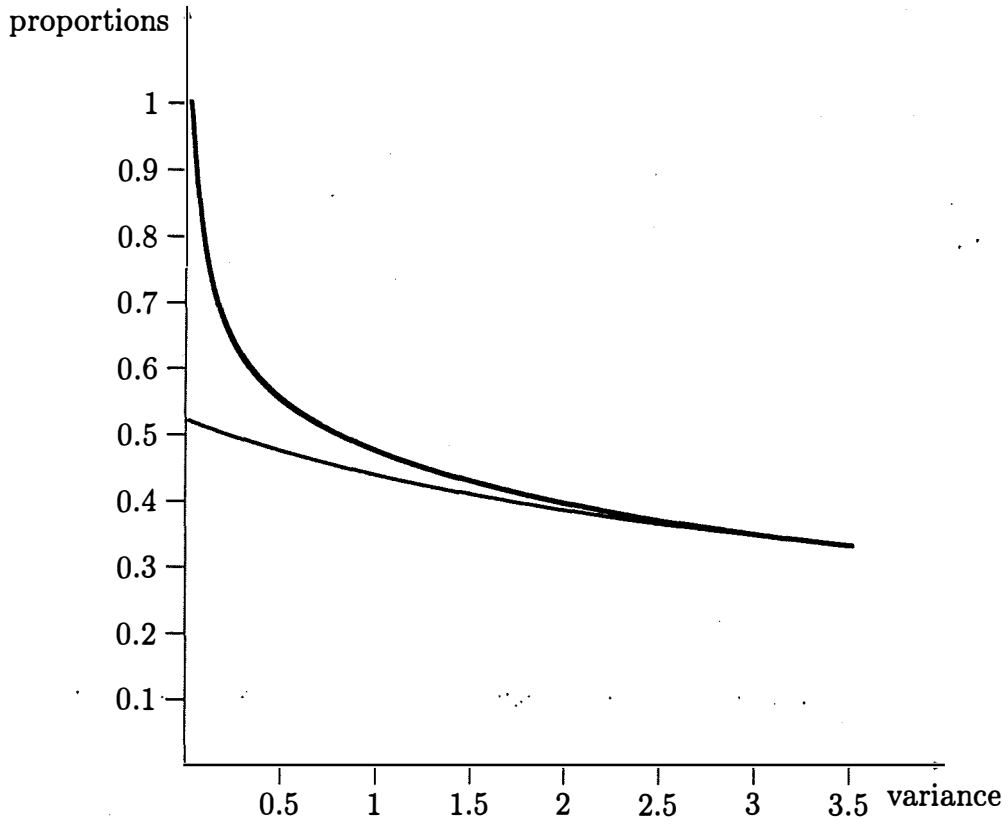
$$\Phi\left(\frac{\widetilde{m}(\sigma_d)}{\sigma_d}\right) - \Phi\left(-\frac{\widetilde{m}(\sigma_d)}{\sigma_d}\right) = 2\Phi\left(\frac{\widetilde{m}(\sigma_d)}{\sigma_d}\right) - 1.$$

From the discussion in 2.3.1, the percentage of voters in the aggregate who approve of centralization is

$$2\Phi\left(\frac{1}{\sqrt{2+\sigma_d^2}}\right) - 1.$$

Therefore the proportion of districts that will approve centralization is greater than the proportion of voters who will if

Figure 1: Proportions of districts and voters voting for centralization as a function of σ_d^2, under sophisticated voting with a continuum of districts The thicker curve indicates the proportion of districts, the thinner line the proportion of voters. They intersect for $\sigma_d^2 \approx 3.032$. The curves stay very close to each other for large σ_d^2. When it is equal to 100, the proportion of districts is 7.38%, the proportion of agents 7.89%.



$$\widetilde{m}(\sigma_d) \geq \frac{\sigma_d}{\sqrt{2+\sigma_d^2}},$$

which, from the definition of $\widetilde{m}(\sigma_d)$ is equivalent to

$$\Phi\left(t^*(\sigma_d) - \frac{\sigma_d}{\sqrt{2+\sigma_d^2}}\right) + \Phi\left(t^*(\sigma_d) + \frac{\sigma_d}{\sqrt{2+\sigma_d^2}}\right) \geq 3/2.$$

Numerical computations show that this inequality holds if $\sigma_d^2 \leq 3$. Figure 1 shows the proportion of districts and of voters that approve centralization, as a function of σ_d^2.

3.2 Optional Confederations

Up to this point, we have assumed that the centralized policy, if adopted, is the median ideal point of all individuals in all districts. This carries with it the implicit assumption that all districts are forced to adopt the same policy and that this policy is somehow decided by majority rule applied to the union as a whole. A natural alternative would be that the centralized policy would be binding only on those districts who have ratified centralization, and that the others adopt the policy preferred by their median voters. A similar "opt-out" rule was used to decide membership in the Confederation of Independent States of the former Soviet Union. Another version was also used for ratification of the U.S. constitution, where nine of the thirteen original states were needed for ratification to be successful, and any state voting against ratification was excluded from the union.[6]

Using the framework with many districts introduced in section 2.3.1, we present three different models of voting behavior, which differ in the degree of sophistication that is assumed on the part of voters, degree of sophistication on which the equilibrium solutions depends crucially. For each of these models, we identify two intervals centered around 0, one is the interval of types of voters who will vote for centralization, the second contains the median of district types for those districts where a majority of voters choose centralization. Because these intervals are centered around 0 the policy chosen by the districts who join the confederation will be 0. Of course, each district that does not join will choose the policy preferred by its median voter.

naive voters: A voter (t, d) votes for centralization if

$$\mathrm{E}\left[(t - m_d)^2 | t\right] \geq t^2.$$

The right hand side is the disutility of the voter under centralization; the left hand side is the expected disutility under decentralization, when the voter does not take into account the fact that his district will vote for decentralization only if the median of its types is far from 0.

It is easy to see that the voters are making exactly the same reasoning that in 3.1, and the analysis can be adapted.

semi naive voters: Districts whose district median lies outside the interval $[-m^{sn}, m^{sn}]$ stay out of the confederation. A voter (t, d) votes for centralization if the expected value of $(t - m_d)^2$ conditional on the fact that he is of type t and on the fact that the inequality

$$|m^{sn}| \geq |m_d|$$

[6] In contrast, ratification of the constitution of the Fifth Republic of France required a majority of voters in a nationwide referendum, as in the analysis of section 2.

Nearly universally, amendment ratification does not have an opt-out provision. For example, in the U.S. ratification of amendments requires appoval by two thirds of the states in the union, rather than some majority or supramajority of voters in a national referendum. These amendments are binding on all states. The exact procedures vary across countries.

holds is greater than t^2, his utility under centralization.

For the sake of brevity, we will not study this case in detail.

sophisticated voters: These voters only take into account in their choice of votes the cases where they will be pivotal. We analyze this case in detail below.

To study tise equilibrium with sophisticated voters, we look for an equilibrium with the following characteristics[7]:

- there exits a cutoff type $\bar{t} \geq 0$ such that a voter votes for centralization if and only if his ideal policy lies in the interval $[-\bar{t}, +\bar{t}]$.
- there exists a cutoff district median $\bar{d} \geq 0$ such that the majority of agents of a district vote for decentralization if and only if the median of their ideal policies lie in the interval $[-\bar{d}, +\bar{d}]$. Therefore exactly 50 percent of the voters of the district with median $\bar{d}$ will choose centralization.
- individuals choose their votes conditional on the fact that they belong to a one of the boundary districts, that is a district with median $\bar{d}$ or $-\bar{d}$. Indeed, this is the only case in which their vote can make a difference.

Because exactly half the agents who live in a marginal district vote for centralization, by a reasoning similar to the one that lead to equation (5), we must have

$$\Phi(\bar{t} - \bar{d}) + \Phi(\bar{t} + \bar{d}) = 3/2. \tag{6}$$

The second condition that we must derive states that an agent of type $\bar{t}$ is indifferent between voting for centralization and decentralization. His vote will make a difference if and only if the median of his district is either $\bar{d}$ or $-\bar{d}$. Let β be the probability that he assigns to the fact that it is $\bar{d}$ (we will compute β later). We must have

$$\bar{t}^2 = \beta(\bar{d} - \bar{t})^2 + (1 - \beta)(\bar{d} + \bar{t})^2. \tag{7}$$

Because of the symmetry of the equilibrium, the districts that belong to the confederation will choose policy 0, and therefore the left hand side represents the utility of the agent if

[7] We ignore deliberately some difficult issues. For example, there clearly exist multiple equilibria: unanimous votes either for centralization or decentralization are clearly equilibria. Also, the identification of a "pivotal voter" in the context of a model with a continuum of voters is questionable. For the purposes of this paper, we feel that the definition that we present is intuitive and enables us to discuss some interesting issues. A complete analysis with a finite number of voters is beyond the scope of this paper.

his district joins it. The right hand side represents his disutility derived from a vote for decentralization. Equation (7) implies

$$\bar{t} = \frac{\bar{d}}{2(2\beta - 1)}. \tag{8}$$

We are left with the task of computing β, that is the conditional probability of belonging to the district of median $\bar{d}$ given that the agent is of type $\bar{t}$ and belongs either to a district of median $\bar{d}$ or to a district of median $-\bar{d}$. Given the distribution of d conditional on t described in 2.2, we get

$$\begin{aligned} \beta &= \frac{\phi\left(\frac{\sqrt{\sigma_d^2+1}}{\sigma_d^2}(\bar{d} - \frac{\sigma_d^2}{1+\sigma_d^2}t)\right)}{\phi\left(\sqrt{\frac{\sigma_d^2+1}{\sigma_d^2}}\,(\bar{d} - \frac{\sigma_d^2}{1+\sigma_d^2}t)\right) + \phi\left(\sqrt{\frac{\sigma_d^2+1}{\sigma_d^2}}\,(-\bar{d} - \frac{\sigma_d^2}{1+\sigma_d^2}t)\right)} \\ &= \frac{e^{\bar{d}\bar{t}}}{e^{\bar{d}\bar{t}} + e^{-\bar{d}\bar{t}}}, \end{aligned} \tag{9}$$

where ϕ is the density of the unit normal distribution. The simplification that yields equation (9) is due to the exponential nature of ϕ.

Solving the system of equations (6), (8) and (9) we obtain

$$\begin{aligned} \bar{t} &\approx 0.714 \\ \bar{d} &\approx 0.339. \end{aligned}$$

Note that $\bar{t}$ and $\bar{d}$ are independent of σ_d^2. This imply that the proportion of voters and of districts that vote for centralization decrease when σ_d^2 increases.

4 Extensions

In this section, we consider two extensions of our basic model.

4.1 Multiple independent issues

Often, centralization or confederation involves loss of sovereignty by each member-district in several policy domains. This is true, for example in the case of the Maastrict agreement, and is obviously true in national confederations such as the U.S. and Canada.

Assume therefore that the constitutional vote bears on the appropriate level of centralization for a very large number K of issues. An individual has a vector of ideal points $t = (t_1, \ldots, t_K)$. If the vector of policies is $x = (x_1, \ldots, x_K)$ his utility will be

$$U(x;t) = -\frac{1}{K}\sum_{k=1}^{K}(x_k - t_k)^2.$$

We assume that the ideal points over different issues are distributed independently[8] from each other by a process identical to that described in section 2.

Since the distribution of ideal points is radially symmetric, a global median exists, and therefore majority rule outcomes on issues are well defined. For an agent with ideal point t the disutility of centralization is

$$\frac{1}{K}\sum_{k=1}^{K} t_k^2 .$$

The disutility of decentralization is

$$\frac{1}{K}\sum \mathrm{E}\left[(d_k - t_k)^2 | t_k\right] = \frac{\sigma_d^2}{\sigma_d^2+1} + \frac{1}{(\sigma_d^2+1)^2}\sum_{k=1}^{K}\frac{t_k^2}{K},$$

where we write d_k for the median ideal point on issue k in the district of the agent.

By the law of large numbers as K becomes very large the expected disutility of centralization converges to σ_d^2+1 and the expected disutility of decentralization converges to 1. *In the limit, as the number of issues become very large, decentralization is preferred by every agent.* A practical requirement of unification is to avoid bundling too many disparate issues together, or perhaps to decide sequentially which issues should have centralized policies.

On the other hand, issues over which one would expect ideal points to be highly correlated (such as agricultural policies for different crops or products) could be bundled without making ratification of centralized policies impossible. Alternatively, to the extent that a few key policy dimensions permit a close approximation of policy preferences over a multitude of specific issues (as is often claimed by political scientists), a large number of issues would not automatically imply unanimous opposition to centralization.

4.2 Uncertainty about preferences

Up to this point, we have assumed that the agents know perfectly their own ideal point. In reality, votes over federation take place when the agents have imperfect information about

[8] At the other extreme, if ideal points are perfectly correlated across issues then nothing changes from the analysis of the last section.

the policies or even the issues to be decided in the future. In order to study this issue, we will use the infinite number of equal size (as in 2.3.1 with $N = +\infty$), and assume that the voters observe their ideal point with a Normally distributed error ϵ. More precisely the agent (t, d) observes

$$I = t + \epsilon,$$

where ϵ is normally distributed with mean 0 and variance σ_ϵ^2.

The variables (d, t, I) are jointly normally distributed with mean $(0, 0, 0)$ and variance-covariance matrix

$$\begin{bmatrix} \sigma_d^2 & \sigma_d^2 & \sigma_d^2 \\ \sigma_d^2 & \sigma_d^2 + 1 & \sigma_d^2 + 1 \\ \sigma_d^2 & \sigma_d^2 + 1 & \sigma_d^2 + \sigma_\epsilon^2 + 1 \end{bmatrix}.$$

Using once again the results discussed in De Groot, we find that conditional on I, the variables (m_d, t) are jointly normally distributed with mean

$$\begin{bmatrix} \dfrac{\sigma_d^2}{\sigma_d^2 + \sigma_\epsilon^2 + 1} I \\ \dfrac{\sigma_d^2 + 1}{\sigma_d^2 + \sigma_\epsilon^2 + 1} I \end{bmatrix}$$

and variance

$$\frac{1}{\sigma_d^2 + \sigma_\epsilon^2 + 1} \begin{bmatrix} \sigma_d^2(\sigma_\epsilon^2 + 1) & \sigma_d^2 \sigma_\epsilon^2 \\ \sigma_d^2 \sigma_\epsilon^2 & \sigma_\epsilon^2(\sigma_d^2 + 1) \end{bmatrix}.$$

We obtain[9]

$$\begin{aligned} \mathrm{E}\left[(m_d - t)^2 | I\right] &= \mathrm{var}\left[m_d | I\right] + \mathrm{var}\left[t | I\right] + \left(\mathrm{E}\left[m_d | I\right] - \mathrm{E}\left[t | I\right]\right)^2 \qquad (10) \\ &\quad -2\mathrm{E}\left[(m_d - \mathrm{E}\left[m_d | I\right])(t - \mathrm{E}\left[t | I\right])| I\right] \\ &= \frac{\sigma_d^2 + \sigma_\epsilon^2}{\sigma_d^2 + \sigma_\epsilon^2 + 1} + \frac{I^2}{(\sigma_d^2 + \sigma_\epsilon^2 + 1)^2}, \qquad (11) \end{aligned}$$

and similarly

$$\mathrm{E}\left[t^2 | I\right] = \frac{\sigma_\epsilon^2(\sigma_d^2 + 1)}{\sigma_d^2 + \sigma_\epsilon^2 + 1} + \frac{(\sigma_d^2 + 1)^2}{(\sigma_d^2 + \sigma_\epsilon^2 + 1)^2} I^2.$$

[9]We do not condition on the name of the district d because all districts are similar.

It is then easy to show that a voter prefers decentralization if and only if

$$\frac{I^2}{\sigma_I^2} \geq \frac{(1-\sigma_\epsilon^2)}{2+\sigma_d^2}.$$

This implies the following:

Proposition 3 *An increase in the uncertainty over preferences, σ_ϵ^2, increases the proportion of voters that prefer decentralization. If σ_ϵ^2 is greater that the variance of types within a district, all agents prefer decentralization.*

The intuition for this result is that on average voters can expect to be closer to their own district median than to the median of the whole. Consider the extreme case, where the variance of the error term is very large so that a voter has no information about his preferences. Then, every voter prefers decentralization since the expected squared distance from his type to the median of his district is less than its expected squared distance from the overall median. This is mathematically equivalent to the many issue case analyzed in 4.2, where the actual distribution of a voter's realized ideal points is asymptotically equal to the ex ante distribution of ideal points.

5 Concluding Remarks

We have explored a model of induced voter preferences over a decision to centralize policymaking. In the uninteresting extreme case (the ex post situation) where voters know everything, the model reduces to the standard one, where the median voter in every district obviously prefers the median of his own district to the median of the whole and centralization cannot arise as a majority rule equilibrium. In the general case, the main theme underlying the derivation of these induced preferences is that voters have incomplete information about preferences of other voters, and therefore view the outcomes of political processes as random variables. However, subsection 4.2 shows that if the voters have no private information about the median of their own district (the ex ante case), centralization never prevails.

But in the broad intermediate range (the interim case), where voters know more about their own preferences than they know about other voters' preferences, while not having perfect information, a much richer set of possibilities emerges. It is possible for a majority of voters in the union of all districts to prefer centralization even though it is common knowledge that the median voter in every district will ex post prefer decentralization. This is true because all voters can expect the centralized policy to be more moderate. We call this the "principle of moderation." An immediate implication is that centralization decisions will inevitably pit the middle against the extremes, with moderate voters favoring centralization and extremists on both wings favoring decentralization. We also

show that it is even easier for a majority of voters in a majority of districts to prefer centralization, which suggests that the actual procedures of aggregating preferences at different levels, perhaps through regional representative institutions like legislatures and parliaments, are likely to have systematic effects in a movement toward more centralized government policies. We call this the "principle of aggregation."

Our model could be extended in a number of directions. One obvious direction is to extend the results to alternative aggregation procedures for deciding over centralized policies. We considered majority rule procedures, other possibilities exist.

Another extension would be to investigate dynamic issues. Just as is it possible for a collection of districts to centralize policies, it is also possible for the reverse to happen. As information about policies and district preferences becomes better known over time, the median voter of each district will have incentives to organize a "secession" of that district from the centralized policies. Thus, it may make sense to have asymmetries (such as supramajority requirements to amend a constitution) built into procedures to reverse earlier decisions to centralize. Without some such devices to "commit" to a centralized policy (including the possible use of force by a central authority), it would be very difficult to avoid the problem of secession, and decisions to centralize might be quite unstable.

References

1. Crémer, J., A. Eustache, P. Seabright, 1993, "Lessons of the Theory of the Firm for the Decentralization of Public Services?" typescript, IDEI, University of Toulouse.

2. de Groot, M., 1970, *Optimal Statistical Decisions*, New York: McGraw-Hill.

3. Oates, W. E., 1972, *Fiscal Federalism*, New York: Harcourt, Brace, Jovanovich

4. Sinn, H.-W., 1993, "How Much Europe? Subsidiary, Centralization, and Fiscal Competition," CEPR *Discussion Paper #834*, Centre for Economic Policy Research, London.